DOCUMENT RESUME

ED 340 761 TM 018 017

AUTHOR Nummedal, Susan G.
TITLE Designing a Process To Assess Higher Order Thinking Skills in College Graduates: Issues of Concern.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
PUB DATE Nov 91
NOTE 33p.; Commissioned paper prepared for a workshop on Assessing Higher Order Thinking & Communication Skills in College Graduates (Washington, DC, November 17-19, 1991), in support of National Education Goal V, Objective 5. For other workshop papers, see TM 018 009-024.
PUB TYPE Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Measurement; *College Graduates; Communication Skills; Critical Thinking; *Educational Assessment; Educational Improvement; *Evaluation Methods; Higher Education; Social Influences; *Student Evaluation; Test Construction; *Thinking Skills
IDENTIFIERS *Higher Order Learning; *National Education Goals 1990

ABSTRACT

The development of a process to assess higher order thinking skills associated with thinking in real world situations raises questions about the domain of assessment, the skills and dispositions to be assessed, and the purpose of the assessment. It is argued that existing critical thinking courses and instruments may be of limited value in designing the assessment. Attention should be given to an assessment of the skills and dispositions of practical intelligence, and assessment measures should incorporate the key characteristics of practical thought, including the social nature. The process of designing the assessment should be a dynamic one, being informed by as many participants representing as many areas of expertise as possible. The assessment process should be seen as a way to further the dialogue among these experts for the goal of improving teaching and learning. A 31-item list of references is included. Reviews by J. Chaffee, P. A. Facione, and R. K. Hambleton of this position paper are provided. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED340761

Designing A Process to Assess Higher Order Thinking Skills in College Graduates: Issues of Concern

Susan G. Nummedal, California State University, Long Beach

Abstract

The development of a process to assess higher order thinking skills associated with thinking in real world situations raises questions about the domain of assessment, the skills and dispositions to be assessed, and the purpose of the assessment. It is argued that existing critical thinking courses and instruments may be of limited value in designing the assessment. Attention should be given to an assessment of the skills and dispositions of practical intelligence, and assessment measures should incorporate the key characteristics of practical thought, including the social nature. The process of designing the assessment should be a dynamic one, being informed by as many participants representing as many areas of expertise as possible. The assessment process should be seen as a way to further the dialogue among these experts for the goal of improving teaching and learning.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Paper prepared for the National Center for Education Statistics Workshop on Assessing Higher Order Thinking and Communication Skills in College Graduates, November 17-19, 1991, Washington, D. C

BEST COPY AVAILABLE

TM018017

Introduction. "The Case of the Ill-structured Problem"

We are all participants in a workshop to develop a process to assess the objective for higher education contained in National Education Goal 5:

The goal:

> "By the year 2000, every adult American will be literate and will possess the knowledge and skills necessary to compete in a global economy and exercise the rights and responsibilities of citizenship."

The objective:

> "The proportion of college graduates who demonstrate an advanced ability to think critically, communicate effectively and solve problems will increase substantially."

The first order of business would seem to be to determine just how this objective relates to the overall goal. The final version of the Author's Briefing Paper (Corrallo, August 12, 1991) links the objective and goal by describing an assessment focusing on higher order thinking skills that (a) would or should be affected by student cognitive and/or affective learning experiences, and (b) would contribute to the development of a competitive workforce and a more effective citizenry. Does this mean that the assessment should address **both** the higher order thinking skills that are or should be acquired through the college experience and those associated with real-world success **or only** those higher order thinking skills acquired through the college experience that are associated with success in the workforce and effective citizenship? If the answer is the former, then it seems reasonable to assume that we may be talking about the assessment of different, thought not mutually exclusive, sets of skills. The exact nature of the overlap and relationship among these sets of skills is far from clear. If the purpose is to assess only those skills acquired through college learning experiences that are associated with success in the workforce and in the exercise of the rights and responsibilities of citizenship, then this is a different task. What we are about remains unclear. Hence, the first issue to be discussed below is this issue of the domain of assessment.

Another question involves the kinds of higher order thinking skills and dispositions that should be included in the assessment. Should they be those that **are being taught currently** in colleges or those that **should be taught** in colleges because they lead to successful functioning in academic settings and/or real world settings? Again, the relationship between the assessment to be designed and these possibilities is unclear. The second issue to be discussed will focus on how to select the skills and dispositions for inclusion in the assessment.

Also unclear is the purpose for which the assessment is to be conducted. Is it to improve instruction? Is it for accountabibility? If so, to whom? For what? The issue of purpose (i.e., assessment for what?) is the last broad issue that will be discussed.

All of these questions serve to make one thing very clear. The task of writing this paper much in common with the most challenging tasks that assess higher order thinking skills: ***it is an ill-structured, "messy" task!*** Hopefully, what follows will give some structure to the process of designing an assessment of the higher order thinking skills of college graduates.

Issue 1. Domain of the Assessment

In trying to determine the relationship between the objective for higher education and the overall goal of adult literacy as described in *America 2000*, I began by raising an issue about the domain of the assessment: Is this an assessment of **both** the higher order thinking skills acquired through college learning experiences and those required for being an effective citizen and competent member of the workforce **or only** those acquired through college learning experiences that are associated with being an effective citizen and competent member of the workforce. I want to argue that we should focus our attention on the latter, namely those skills acquired through learning experiences in college that are relevant to successful functioning in real-life situations.

It may be helpful to remember that when specific courses in critical thinking have been introduced into the curriculum, one of their major goals has been to enable students to think more effectively about everyday issues and concerns, such as practical problem solving and decision

making. Indeed, some (e.g., Sternberg, 1987) have even argued that this is the most important goal of such a course. This goal has been described variously as the goal of "transfer," "generalization," or "application." Concern with this goal has led, for example, to the controversy over the relative benefits of teaching general reasoning processes vs. domain specific processes (Glaser, 1984). It is inherent in questions such as: "How can logic be made more relevant to everyday reasoning?" (Baron, 1988, p. 153) and "What can one do to maximize transfer of training for thinking skills to students' everyday lives?" (Sternberg, 1987, p. 258).

Notice that the starting point in these questions seems to be the desire to find better ways to teach those skills that have already been identified as core to thinking critically such that successful "transfer" or "generalization" or "application" to problems of everyday thinking can take place. This may be the wrong starting point. Rather that asking, "How can those skills we have identified for teaching be made more relevant to everyday reasoning?" or "What teaching strategies can be employed to insure transfer of cognitive skills to everyday reasoning contexts?" we may need to start with a different question: "What is the nature of everyday thinking?"

Applying this distinction to the purposes at hand, I believe we need to start with this same question about the nature of everyday thinking rather than with the question about how the skills identified for college instruction should be manifested in the everyday thinking associated with effective citizenship and success in the workforce. (This issue will also be raised below in the context of the discussion of issues associated with selection of skills and dispositions)

What is the nature of everyday thinking, practical thought, and practical problem solving? To begin to answer the question, it is important to distinguish between what Neisser (1976) has called academic intelligence (i.e., what it takes to be successful in school settings) and practical intelligence (i.e., what it takes to perform successfully in real world settings). Others have characterized the dichotomy as the difference between theoretical and practical intelligence (Scribner & Fahnmeier, 1982), academic and practical intelligence (Ford, 1986), and school success and life success (McClelland, 1980). According to Neisser (1976), tasks typically found in school

settings are formulated by other people and most often provide complete information, offer no additional information, feedback, or opportunities for growth in thinking during the solution process, and require an unemotional and conflict-free approach for successful solution. In contrast, he characterizes practical intelligence as "passionate, emotional, influenced by internal states that have both long- and short-run dynamics of their own" (p. 137). Tasks embedded in real world settings offer opportunities for satisfying both practical and interpersonal motives simultaneously. Prior knowledge, beliefs, and attitudes as well as specific skills are all brought into play in working with problems in everyday settings. To the extent that (a) "practical intelligence" is requisite for competing in a "global economy" and "exercising the rights and responsibilities of citizenship," and (b) we measure "life success" in part by being able to do these things effectively, then we need to know what these ingredients of practical intelligence and life success are.

Scribner (1986) describes practical thought as "thinking in action" or "thinking that is embedded in the larger purposive activities of daily life and that functions to achieve the goals of those activities" (p. 15). Similarly, Ford (1986) defines practical intelligence as "the attainment of relevant goals in specified environments, using appropriate means and resulting in positive developmental outcomes" (p. 183).

In studies of practical, "on the job" thinking in a variety of occupational settings, several investigations reveal that real-world problem solving can be characterized by a number of common features. That is, both the problems themselves and the practical thinking necessary to deal successfully with them share a number of common feature.

First, with regard to the nature of the problems themselves, everyday problems are quite often ill-structured or "messy" problems. They tend to be problems that don't look exactly like things we're seen before. Even though they may appear to be *formally* very much like previously solved problems, they may indeed be *functionally* unique due to the specific aspects of the context in which they reside. Because of their ill-structured nature, they must be "reconstituted" and given a definite formulation within the existing constraints of the setting in which they

reside. Thus, one feature of everyday problem solving is that it requires formulating or redefining the problem itself as well as generating problem solutions. Goodnow (1986) has presented convincing evidence documenting the importance of organizing and, particularly, *reorganizing*, to intelligent behavior in daily life. This would seem to imply that individual's who are successful in solving real-world problems are able to differentiate between well-structured and ill-structured problems, devoting some of their resources to restructuring and reorganizing where appropriate.

Second, given the ill-structured nature of everyday problems, it should not be surprising to find that two of the most important features characterizing solutions to these problems include efficiency and flexibility. A high priority in everyday thinking seems to be "efficiency" (Rogoff & Lave, 1984), "economy of effort" (Scribner, 1984), "minimizing cognitive load" (Nickerson, Perkins, & Smith, 1985). As Rogoff and Lave (1984) indicate,

> "what is regarded as logical problem-solving in academic settings may not fit with problem solving in everyday situations, not because people are "illogical" but because practical problem-solving requires efficiency rather than a full and systematic consideration of all alternatives. Rather than employing formal approaches to solving problems, people devise satisfactory opportunistic solutions. Everyday thinking, in other words, is not illogical and sloppy but instead is sensible and effective in handling the practical problem" (p. 7).

Flexibility, or the ability to shift among solution strategies as the problem space requires, also has been identified as a key component of practical thinking across a wide variety of occupations, including the "informal improvisation" associated with professional expertise (Schon, 1983). Scribner (1986) has pointed out that effective problem solvers are able to use of the environment (including its social, symbolic, and material resources) in which the problem is situated to effect better problem solutions. Likewise, Goodnow (1986) points out that recognizing constraints (i.e., the incompatible and unalterable components) as well as sources of freedom (i.e., the combinable and malleable components) is also important in everyday functioning. This reinforces what numerous others (e.g., Sternberg, 1982) have identified as the central role of meta-

cognitive processes in directing everyday problem solving.

The use of flexible, least effort strategies has been described by Rogoff and Lave (1984) as necessary, since everyday problems often involve unanticipated opportunities and constraints. "A more systematic and precise approach would result in less effective practical action since it would take more effort to develop and would be less flexible" (p. 7).

And finally, there is a very important **social** component to everyday thinking and problem solving. As Goodnow (1986) notes, "it is very rare in everyday life that organizing and reorganizing take place without other people being involved" (p. 153). The model of the individual problem solver sitting down alone to face a problem and come up with a solution independent of input from other people may be very rare indeed - and certainly cannot be assumed to be the first step in problem solving.

Implications for assessment.

This discussion of everyday thinking points to two important implications for assessment. First, implicit in the characterization of everyday thinking as being flexible and efficient is the notion that people have available multiple strategies for solving problems **and** that the first one(s) selected may often represent something *less than their "highest" level of thinking* about solutions to the problems at hand. If we give a person just "one shot" at solving a problem or addressing an issue (even if this "shot" is under the guise of a ***National Test***) we need to be very careful about what we conclude about what that person knows or does not know. To more closely approximate everyday thinking and problem solving, a performance based assessment needs to be a dynamic process. For example, an assessment might start with a person working with some materials and making some response(s). Feedback as to the adequacy of the response(s) would then be given, either in the form of "trying out" the solution to see how it works or in the form of some "expert" commenting on the adequacy of the solution. There could be several iterations of this process until the person thought the problem was "solved."

This notion of feedback is directly tied to a second point noted above, namely, that everyday thinking and problem solving most often occurs in interaction with others. Thus, one type of assessment that

needs serious consideration is an assessment within a group context. Asking students to solve problems in conjunction with others would more nearly model real world problem solving. Since this assessment is not designed to provide information about *individuals*, there is no reason to reject outright a group problem solving exercise. Of course, there are many other grounds on which to reject such a suggestion, but before doing so, please read on. Quite obviously, there are substantial problems with the lack of control over the characteristics of individuals and the nature of their input in a group problem solving situation. But some of these concerns could be dealt with by allowing group interaction while still requiring each individual to respond to a series of specific questions. Or each group member could be assigned a role to play when addressing the problem and asked to solve the problem from the perspective of that role. Perhaps the nature of the group interaction could be structured to cover particular specified points. Over the course of the problem solving interaction, individual members playing specific roles could respond to a set of questions specific to that role. Responses could then be aggregated across roles.

An interactive computer dialogue also could be designed where the individual student is part of a simulated group, getting input as well as feedback from other "members" of the group. HyperCard would seedm to be partricularly well suited for this purpose. In this case even more control over the nature of the interactive process could be exercised once the prototype had been developed.

Issue 2. Selection of Skills and Dispositions for the Assessment

How can we use information about what is being taught in courses in critical thinking to guide the selection of skills and disposition in this assessment? If we look at the higher order thinking skills that are being taught currently in college, my sense is that, with some notable exceptions (e.g., Alverno College), we are looking at discipline specific conceptualizations of critical thinking and problem solving. Even on campuses where specific coursework in critical thinking has been mandated (e.g., campuses in the California State University System), courses meeting the requirement are taught in different departments

(most often philosophy and psychology). These departments have developed courses which derive from specific frameworks of critical thinking tied to theoretical formulations grounded in a specific disciplines.

Kurfiss (1988) has described three of the dominant approaches to critical thinking instruction. The traditional approach, emphasizing argument analysis and informal logic, is taught in the humanities, particularly philosophy and English. A second approach, emphasizing information processing and judgment skills such as problem solving, decision making, and scientific reasoning, is most often taught in the social and behavioral sciences, particularly psychology. Both of these approaches are similar in that they stress specific skills which are assumed to be teachable, although the particular skills targeted for instruction differ. The third approach derives from cognitive developmental theory. In courses designed around Piaget's theory of cognitive development, instruction focuses on the skills associated with formal operational thinking (Inhelder and Piaget, 1958). Another cognitive development approach is based on Perry's work (1970) and stresses the student's espitemological world view (i.e., beliefs about the nature and source of knowledge). Rather than being the product of a specific course or particular series of courses, the development of this world view is thought to be the product of the total college experience both inside and outside the classroom. In this approach, the focus is primarily on dispositions, as opposed to specific skills, associated with critical thinking.

The notion of explicit instruction in critical thinking has been the subject of continuing debate. Much of this debate is over the issue of whether critical thinking is discipline specific or can be taught generically. This debate is far from over.

We still know relatively little about the effectiveness of various college courses in meeting their specific instructional goals. We know even less about the *relative* effectiveness of these courses in meeting the more general goal of improving the thinking abilities of students. One notable exception is the recent research by Lehman and Nisbett (1990) demonstrating that specific kinds of thinking skills are differentially enhanced as a function of particular kinds of educational experiences.

I believe we need to know more about these questions of instructional effectiveness. The knowledge gained will help in both shaping the current assessment and interpreting its results. But I do not believe that an analysis of what is being taught currently will by itself be sufficient. At most it will be suggestive of some of the skills and dispositions that might be considered for inclusion. Too many questions remain as to how these skills relate to performance in the real world situations to consider such an analysis an end in itself.

How can we use information about what is currently contained in various measures of critical thinking to guide the selection of skills and dispositions? Here, I think we already know a great deal about the shortcomings of some of the most widely used tests of critical thinking. Those who have been engaged in the business of assessment at the state level (e.g., New Jersey) have already been through a process of examining various "off the shelf" measures such as the Watson-Glaser Critical Thinking Appraisal (e.g., Watson & Glaser, 1980) and have found them wanting. Morante (1984), for example, reports on the problems of *cleanness* (i.e., the confounding of reasoning with reading abilities) associated with the Watson-Glaser. And Norris (1988a and 1988b), in an analysis of items from the Watson-Glaser and the Cornell Critical Thinking Test, Level X (Ennis & Millman, 1985), argues that "there is no reason to believe that, *in general*, when examinees choose keyed responses they think critically and when they choose unkeyed responses they think uncritically" (p. 7). Reports such as these warn us about the difficulty of inferring higher order thinking from performance on these measures. While these tests per se may be substantially flawed as measures of critical thinking, the question still remains as to whether the skills and dispositions they purport to measure might not be appropriately selected for inclusion in this assessment.

Implications for assessment.

For the most part, I have the same reservations about these tests as I have about existing critical thinking courses as guides for selecting the skills and dispositions to be included in this assessment. I think in both cases, the skills and dispositions are derived from conceptions of critical

thinking that are both narrow and discipline specific. Rather than starting with existing courses and tests, we need to look to the world of practical thinking and identify those skills and dispositions associated with successful performance. I believe we should build on the experiences of those who already have been about the business of creating authentic, performance-based measures (Archbald & Newmann, 1988), examining these measures for higher order thinking skills and dispositions that appear to be associated with performance in real world situations. (I am aware of the irony of, on the one hand, seriously questioning existing instruction as it relates to success in real-life situations, and, on the other hand, going about an assessment that is designed "precisely" to determine how well colleges are doing in teaching the skills and dispositions necessary for success in real-life situations.)

A further point about authentic assessments needs elaboration. Performance assessments typically consist of far fewer tasks, each one of which takes far longer to perform. (For what is perhaps a "record" for sustained intellectual performance, see the French and German national tests (Cheney, 1991) that allow four hours to complete a single essay response). These are assessments that are designed to "go deep," with beliefs, prior knowledge, and experience being more likely to play an important role in performance. And, if an assessment does "go deep," then practical limitations make it difficult to also "go broad" (i.e., have many tasks which sample a variety of content areas. If indeed, probing deeply into the thinking of students is desireable, it may be important, as Linn (1991) has suggested, to include many "deep" tasks representing a broad array of content areas for a particular skill and allow examinees to choose among them.

It should also be noted that no single skill or disposition can be equated with higher order thinking. True, individual skills may be activated in the service of some higher order thinking activity. But when one engages in an activity requiring higher order thinking, a unique subset of skills will be invoked in the service of that activity and may be used and combined in ways unique to that particular activity. While the individual skills form the "building blocks" of performance, it is incorrect to assume, as Pottinger (1978) has noted, that the mere summation of performance on objectively measurable and highly specific subtasks can

be equated with higher order thinking competence. As a result, authentic, performance-based tasks should be designed which explicitly incorporate a number of specific skills. Given the complex nature of these tasks, performance on one task may be judged for several purposes corresponding to different skill assessments.

We also need a place for self assessment in the assessment process. If we take seriously the goal of becoming a "nation of students," if we recognize that learning is a life-long process, then we must be certain to address the metacognitive skills associated with self-monitoring and self-assessment. Inclusion of measures of self-evaluation, such as asking students to rate their own performance on the assessment tasks, should be considered.

Issue 3. Assessment to Inform.

When we talk about an assessment designed to inform practice, one question we must ask ourselves is, "Where is instruction in critical thinking, problem solving and communication that we want to inform taking place?" There are several possibilities, including (a) in separate courses explicitly designed to teach these attitudes and skills (e.g., in courses explicitly dedicated to instruction in critical thinking, written communication, etc.); (b) within courses where instruction in these skills is just one of several explicit goals and not the most important one at that (e.g., some universities require some minimum amount of instruction in critical thinking as an explicit goal in each general education course); (c) within courses where instruction in one or more of these skills is not an explicit goal but may be thought to be a by-product of instruction - indeed even may be requisite for success in these courses - but where no specific instruction in these skills takes place); (d) within the major area of study (e.g., in required research methods courses); (e) as a product of the general university experience (e.g., Perry, 1970). These possibilities, of course, are not mutually exclusive. To the extent they exist in various combinations across different institutions, it becomes difficult to talk about "informing instruction" as if instruction were a unitary thing. At a minimum, differences in practices associated with instruction in these skills must be recognized and, where possible, documented.

Implications for Assessment.

Perhaps one way to approach the task of informing instruction would be to develop a typology of institutions based upon where the goals of instruction in critical thinking, problem solving, and communication are thought to be realized in the curriculum or even a quantitative measure that rates the proportion of students' instruction that is devoted to these goals. Responses could then be aggregated within types of institutions based on pre-established institutional types. But deciding where in the curriculum and the degree to which these goals are being realized is no easy task. What are we to take as evidence for the claim that these goals are being addressed? Content analyses of college catalogues, general education guidelines, and other official documents might prove useful. These, however, are relatively distal measures of these goals. More proximal measures might be derived by examining course syllabi, assignments, and examinations used in specific courses. Or one could analyze student products for evidence of the implementation of these goals (i.e., portfolio assessments). One could further refine this approach by assessing the educational context for individual students based upon course patterns within types of institutions.

While this approach would be woefully lacking in providing information on student performance to individual instructors of different "kinds" of courses, it would provide important information on the performance of students from different educational contexts within different institutional types. In a recent article on kinds of intelligence, Ceci and Liker (1986) argue persuasively for a contextual account of intelligence, concluding that "persons develop in context and it is important to know something about their developmental contexts before we pass judgment on their intelligence" (p. 138). To extend the argument to the present context, we might say, "college students are educated in context, and it is important to know something about their educational contexts before we pass judgment on their higher order thinking skills."

But this assessment is not just about informing instruction. It is about informing practice and practitioners of all stripes - students, educators, employers, assessors, and policy makers. If we see the

primary purpose of this assessment as ultimately improving teaching and learning, then we will view it as the start of a dialogue **with**, not a report **to**. We will see the assessment as a dynamic process which engages all the players. For example, college instructors need to not only be informed of results but inform the assessment process. Their expertise needs to be drawn upon to help shape the assessment tasks and guide the process of setting performance standards. So too for other professionals. We will see it as an ongoing process. There are far too many questions to which we have no answers. We are far from a full understanding of the nature of higher order thinking skills. We are far from understanding how the educational process might bring about the kinds of changes in higher order thinking we are trying to assess. We might do well to think of this assessment as an opportunity to discover what some of the interesting questions about this process really are rather than as a final report on "what is."

Other Important Issues to Consider.

Several other important issues to consider in designing the assessment process are discussed below. Each is important in its own right and worthy of further exploration. I raise them in the hope that they will be considered more fully in the workshop.

1. We simply must have baseline data for the skills and dispositions we choose to assess. We need to know where the students start with respect to these specific skills and dispositions in order to make sense out of the data describing where they end up. Ideally, the assessment should have a longitudinal research design with the same individuals measured over the course of the college years. Given the difficulties associated with such a design, there should at least be a cohort design to provide a picture of the entry level higher order thinking skills and the changes that occur over the college years.

2. We need to be prepared for two kinds of increases associated with education: increases in average level of performance and increases in variability. Some institutions that have been tracking their students over

the college experience (e.g., Evergreen College) report both kinds of increases. The objective for higher education appears to address average performance levels only. Do we want to make any predictions about increases in variability and what they might mean? How might increases in variability relate to different skill and disposition patterns across students?

3. At the Author's Briefing Meeting, there was some discussion about the need to consider multiple models for outcomes, recognizing that America needs a diverse workforce with a variety of skills and abilities. In his recent book, *Powershift*, Alvin Toffler talks about the need to shift our metaphor from America the "melting pot" to America the "salad bowl," from a conception of "mass democracy" to one of "mosaic democracy." He argues persuasively for role of diversity in contributing to economic growth. The work of Kaplan (1984) raises the serious issue of culturally specific conceptions of critical thinking, problem solving, and communication. As the assessment is designed, we need to remain conscious of the extent to which both the higher order thinking and communication skills selected and the ways in which they are measured reflect the diversity of experience of college graduates and the diversity of experience needed in an effective workforce. We need to see this diversity as an asset, not a problem to fix, even if this calls into question some of our basis assumptions about the universal nature of our current models of critical thinking and problem solving.

4. We should be aware that the kinds of skills and dispositions necessary to meet the goals of success in the workforce, effective exercise of the rights and responsibilities of citizenship, and life-long learning most likely will not be well developed in the typical college graduate of 22 years of age, not even in the 30 year old graduate. Attributes associated with reflective judgment (i.e., "recognition of the limitations of personal knowledge, acknowledgment of the general uncertainty that characterizes human knowing, and humility about one's own judgments in the fact of such limitations") involve an epistemological stance rarely found in young college graduates (Kitchener & Brenner, p. 126).

5. We need to incorporate into the assessment skills associated with cooperation. The document, *America 2000*, speaks of a competitive workforce. Nowhere does it speak about the importance of a *cooperative* one. Given my value system, I believe we should focus at least as much on measuring the success of our institutions of higher learning in promoting cooperation as in promoting competition.

References

Archbald, D. A., & Newmann, F. M. (1988). Beyond standardized testing: Assessing authentic academic achievement in the secondary school. Reston, VA.: National Association of Secondary School Principals.

Baron, J. (1988). Thinking and deciding. Cambridge: Cambridge University Press.

Ceci, S. J., & Liker, J. (1986). Academic and nonacademic intelligence: An experimental separation. In R. J. Sternberg and R. K. Wagner (Eds.), Practical intelligence: Nature and origins of competence in the everyday world (pp. 119-142). Cambridge: Cambridge University Press.

Cheney, L. V. (1991). National tests: What other countries expect their students to know. Washington, D. C.: National Endowment for the Humanities.

Corrallo, S. (1991, August 12). Author's briefing paper. Washington, D.C.: National Center for Educational Statistics.

Ennis, R. H., & Millman, J. 91985). Cornell Critical Thinking Test, Level X. Pacific Grove, CA: Midwest Publications.

Ford, M. E. (1986). For all practical purposes: Criteria for defining and evaluating practical intelligence. In R. J. Sternberg and R. K. Wagner (Eds.), Practical intelligence: Nature and origins of competence in the everyday world (pp. 184-200). Cambridge: Cambridge University Press.

Glaser, R. (1984). Education and thinking: The role of knowledge. American Psychologist, 39, 93-104.

Goodnow, J. J. (1986). Some lifelong everyday forms of intelligent behavior: Organizing and reorganizing. In R. J. Sternberg and R. K. Wagner (Eds.), Practical intelligence: Nature and origins of competence in the everyday world (pp. 143-162). Cambridge: Cambridge University Press.

Inhelder, B., & Piaget, J. (1958). The growth of logical thinking from childhood through adolescence. New York: Basic.

Kaplan, R. B. (1984). Crultural thought patterns in inter-cultural education. In S. McKay (Ed.), Composing in a second language. Cambridge, MA.: Newbury House.

Kitchener, K. S., & Brenner, H. G. (1990). Wisdom and reflective judgment: Knowing in the fact of uncertainty. In R. J. Sternberg (Ed.), Wisdom: Its nature, origins, and development. Cambridge: Cambridge University Press.

Kurfiss, J. G. (1988). Critical thinking: Theory, research, practice, and possibilities. ASHE-ERIC Higher Education Report No. 2. Washington, D.C.: Association for the Study of Higher Education.

Lehman, D., & Nisbett, R. (1990). A longitudinal study of the effects of undergraduate training on reasoning. Developmental Psychology, 26, 952-960.

Linn, R. (1991, June). Fair enough? A closer look at alternative assessment. Paper presented at the Sixth AAHE Conference on Assessment in Higher Education, San Francisco.

McClelland, D. C. (1980). Motive dispositions. In L. Wheeler (Ed.), Review of personality and social psychology, Vol. 1. Beverley Hills: Sage.

Morante, E. A., & Ulesky, A. (1984). Assessment of reasoning abilities. Educational Leadership, 42, 71-74.

Neisser, U. (1976). General, academic, and artificial intelligence. In L. B. Resnick (Ed.), The nature of intelligence. New York: Wiley and Sons.

Nickerson, R. S., Perkins, D. N., & Smith, E. E. (1985). The teaching of thinking. Hillsdale, N. J.: Erlbaum.

Norris, S. P. (1988a). Controlling for background beliefs when developing multiple-choice critical thinking tests. Educational Measurement: Issues and Practice, 7, 5-11.

Norris, S. P. (1988b). Research needed on critical thinking. Canadian Journal of Education, 13, 125-137.

Perry, W. G., Jr. (1970). Forms of intellectual and ethical development in the college years: A scheme. New York: Holt, Rhinehart, and Winston.

Pottinger, P. S. (1978). Designing instruments to measure competence. Journal of Teacher Education, XXIX, 28-32.

Rogoff, B., & Lave, J. (Eds.) (1984). Everyday cognition: Its development in social context. Cambridge: Harvard University Press.

Schon, D. A. (1983). The reflective practitioner. New York: Basic.

Scribner, S. (1984). Studying working intelligence. In B. Rogoff & J. Lave (Eds.), Everyday cognition: Its development in social context (pp. 9-40). Cambridge: Harvard University Press.

Scribner, S. (1986). Thinking in action: Some characteristics of practical thought. In R. J. Sternberg and R. K. Wagner (Eds.), Practical intelligence: Nature and origins of competence in the everyday world (pp. 13 - 30). Cambridge: Cambridge University Press.

Scribner, S., & Fahnmeier, E. (1982). Practical and theoretical arithmetic: Some preliminary findings. Industrial Literary project Working Paper No. 3. New York: The Graduate School and the University Center, City University of New York.

Sternberg, R. J. (1982). Training of intelligent performance. In D. K. Detterman & R. J. Sternberg (Eds.), How and how much can intelligence be increased (pp. 141-146). Norwood, NJ: Ablex.

Sternberg, R. J. (1987). Questions and answers about the nature and teaching of thinking skills. In J. B. Baron and R. J. Sternberg (Eds.), Teaching thinking skills: Theory and practice (pp. 251 -259. New York: Freeman and Co.

Watson, G. & Glaser, E. M. (1980). Watson-Glaser Critical Thinking Appraisal. Cleveland, OH: Psychological Corporation.

Fiorello H. LaGuardia Community College THE CITY UNIVERSITY OF NEW YORK
31-10 THOMSON AVENUE, LONG ISLAND CITY, N.Y. 11101 • Telephone (718) 482-7200

John Chaffee, Ph.D.
Director, Creative and Critical Thinking Studies
LaGuardia Community College; The City University of New York

Review of the position paper: *Designing A Process to Assess Higher Order Thinking Skills in College Graduates: Issues of Concern*
Susan G. Nummedal, California State University, Long Beach

1. What abilities (critical thinking, communication, problem-solving) have been identified and why were they selected?

The major thrust of this paper is to emphasize the need for focusing on those cognitive and communication abilities that fall into the category of "practical intelligence" used to think effectively in "everyday, real world situations." These abilities are contrasted with those that fall into the category of "theoretical, academic intelligence" used to achieve success in "school situations."

These categories are defined in only the most general way. Following Neisser's (1976) characterization, "theoretical, academic intelligence" addresses tasks which are typically formulated by other people and most often provide complete information, offer no additional information, feedback or opportunities for growth in thinking during the solution process, and require an unemotional and conflict -free approach for successful solution." In contrast, "practical intelligence" is "passionate, emotional," and is used to address tasks in which "prior knowledge, beliefs, and attitudes as well as specific skills are all brought into play." Problems in the "real-world" setting are often "ill-structured" and "messy,"necessitating cognitive "efficiency and flexibility" and social collaboration with others.

In my opinion, the general emphasis of the author on the need to assess cognitive and communication abilities within diverse, practical contexts is appropriate and well-reasoned. However, I have concerns regarding her specific formulation of this project. With the exception the type of general characterizations described above, the crucial thinking abilities to be addressed in this project are not specifically defined. This is due at least in part to to fact that the general categories used are conceptually confused, as I will try to explain.

The dichotomy between "academic" and "real-world" contexts is problematic. I believe that the author is quite right in insisting that a meaningful assessment of cognitive abilities involve practical contexts and real-world situations. However, *effective*, academic instruction does in fact integrate practical, real-world problems, issues, case-studies, and learning situations. This integration certainly needs to be expanded and informed by a broad range of learning contexts, but it is not accurate or helpful to portray the academic arena exclusively as a self-contained, hermetically-sealed, educational biosphere, completely isolated from practical considerations and life experience.

Further, the dichotomy between "theoretical intelligence" and "practical intelligence" is also problematic. The fact is that theoretical understanding and practical application function most effectively in dialectical integration, whether the context is academic or non-academic. To paraphrase Immanuel Kant: "Applications without conceptual understanding is blind; conceptual understanding without application is empty." In order to develop complex cognitive and communication abilities in a meaningful and lasting fashion, people need to develop both theoretical, conceptual frameworks and the ability to apply these frameworks to practical contexts. This should be a guiding principle for the current NCES project.

2. Are the selected abilities appropriate in meeting the goals of this project?

Since the author emphasizes developing and assessing abilities derived from real-life contexts, the abilities she believes to be important are by definition appropriate to the workplace and in the practice of citizenship -- the goals of the project. However, these abilities are are not specified or defined by the author. She is describing a methodological approach rather than articulating actual cognitive and communication abilities.

3. Are the abilities defined in a way that would make possible assessing their development in college students?

Since the author does not define the abilities to be developed and assessed, except in general terms, this question cannot be addressed.

4. Do the proposed assessment methods allow for: accurately measuring the abilities; determining the acquisition barriers; identifying effective learning environments?

The analysis presented in this paper does not include specific methods, and so this question cannot be addressed.

5. Are the methods or suggestions presented practical, replicable and complete.

For the reasons noted above, the content of this paper does not address the issues in this question.

6. General Comments:

Conclusion: This paper presents a general review of the issues and challenges posed by the proposed national assessment of cognitive and communication abilities. It contains useful insights and references, although for the most part the author's analysis does not address the specific assessment issues embodied in the evaluation criteria.

Amendment: The Creative and Critical Thinking Program at LaGuardia Community College:

The Creative and Critical Thinking program at LaGuardia is based on the assumption that thinking is a process that can be understood and improved through proper study and practice. The keystone of this nationally recognized program is *Critical Thinking Skills*, an interdisciplinary course developed in 1979 which explores the cognitive process and helps students develop the higher-order thinking, reasoning, problem-solving and literacy abilities needed for academic and career success. In addition, the course is designed to explore and develop these abilities in the context of personal growth and social awareness, fostering qualities like mature judgment and social responsibility.

This interdisciplinary course has become an integral part of the college's curriculum, and is taught to over 800 students annually by faculty from a wide variety of disciplines. The curriculum for the course is based on the text Thinking Critically (3/E), authored by John Chaffee and published by the Houghton Mifflin company. A second course, entitled *Creative Thinking: Theory and Practice*, focuses on the cognitive processes we use to generate and refine innovative ideas. This course utilizes a unique blend of textbook readings, research articles, guest speakers and creative projects.

The development of the program has been fueled by grants from The National Endowment for the Humanities (NEH), designed to integrate critical thinking abilities across the college's curriculum. The vehicle for curriculum infusion has been course "pairings," in which sections of *Critical Thinking Skills* are linked to courses from other academic areas and accompanied by faculty seminars.

Students involved in the program have consistently demonstrated improved thinking abilities and accelerated development of language skills. The program has been evaluated by The Educational Testing Service as "a mature educational program which has earned a high degree of success with a wide variety of students," and it has been cited by NEH as "a very enlightened approach to undergraduate instruction." As detailed in "The Final Report to NEH" (available by writing to the author), the program appears to have succeeded in meeting its three primary objectives: **literacy, reasoning and problem-solving,** and **critical attitudes.**

(1) **Reasoning and Problem-Solving:** Utilizing a variety of evaluation strategies, the major evaluator of the project, Dr. Garlie Forehand, Director of Research, Program-Planning and Development, at the Educational Testing Service, concluded that the program fosters the development of students' thinking abilities at both general and specific levels.

"At the general level, teachers perceive more respect for the thinking process, more tendency to bring a "habit of thinking" to their classes. At the specific level, teachers reported instances of transfer of such skills as breaking problems into parts, classifying, organization of thought, asking questions, separating facts from opinions, and assessing alternative points of view."

Students also recognized the development and transfer of thinking skills from *Critical Thinking Skills* to other content courses, citing examples like breaking problems into parts in math, applying the concepts of perceiving to the concept of ethnocentrism in social science, transferring self-perception insights to oral communication, and so on. Since fundamental thinking abilities and critical attitudes work together and interact in complex ways, students do not learn them in a skill-by-skill fashion. Instead, concurring with developmental theory and faculty analyses, students in the program seem to be undergoing a developmental process in which skills, attitudes and perceptions are progressively reorganized into new cognitive patterns. This leads to breakthrough or "aha" experiences as students discover new methods and abilities, revealed in student comments like: "It expands thinking -- like a tool"; "Part of my brain awakened"; "It put a seed, a spark, in me."

(2) **Literacy:** Since language and thinking are such closely related, reciprocal and interactive processes, the LaGuardia program is designed to improve students' thinking abilities while simultaneously enhancing their language skills. The cumulative results of the program have revealed that students enrolled in *Critical Thinking Skills* pairs have consistently demonstrated accelerated development of language skills as measured by standard language examinations. In addition to improvements in students' grammatical and structural language skills, faculty also report that students are learning to use language with a depth, insight and sophistication unusual for students at this level, as they seek to utilize and express their evolving higher-order thinking abilities.

(3) **Critical Attitudes:** One of the guiding principles of the Creative and Critical Thinking program is the belief that learning should take place in experiential contexts, serving to stimulate qualities such as self-awareness, initiative and maturity. As Dr. Forehand notes, faculty reported that students displayed this sort of affective development by being more attentive, less likely to be absent, more quick to follow instructions, more serious about course work, better at asking questions, better at verbalizing, less afraid of thinking and expressing themselves, and evidencing increased self-confidence." One professor summed up: "Maybe maturity is the word."

Becoming a critical thinker does not simply involve developing discrete intellectual abilities: in involves developing insight, reflective judgment, informed beliefs, and a willingness to carefully explore diverse perspectives with incisive questions. As students develop their critical thinking abilities, they also grow as individuals, developing the qualities of openmindedness, responsibility, initiative, and a sense that they can control the direction of their lives through the choices that they make. In the words of one LaGuardia student:

"The words 'critical thinking' will never leave my vocabulary because by learning how to organize my ideas, support my point of view with reasons and trying to solve my problems rationally, I have learned more effective ways of dealing with my life, my children and my schoolwork."

In sum, the Creative and Critical Thinking program works a knitting together thinking and literacy abilities with the fabric of students' experience. The program rests on the conviction that thinking and literacy abilities can only be taught effectively through a process of synthesis, giving students the means to clarify and make sense of themselves and the world in which they live.

A Critique of Susan G. Nummendal's

"Designing A Process to Assess Higher Order Thinking Skills in College Graduates"

Prepared by Peter A. Facione

Dr. Nummendal argues that a focus on "life success" in "everyday" situations, which are characterized by "ill-structured" problems calling for "practical intelligence" with its "social component," raises a host of new questions and challenges for critical thinking assessment. Dr. Nummendal says she thinks that virtually all conceptualizations of critical thinking currently operative in college teaching fail because they are too narrow and discipline specific. She reminds us that the most widely used, commercially available CT tests have shortcomings, and she warns that they "may be substantially flawed."

Let's set aside the myriad unresolved theoretical issues regarding various conceptualizations of human intelligence and how these might relate to CT. Let's not take up the host of questions regarding how to operationalize the crucial vagaries of the five terms in quotes in the first sentence. Let's not look too closely at the off-handed and incomplete treatment of the currently available commercial CT tests. Instead, given the purpose of her paper is to revisit some basic issues in CT assessment, let's focus on the plausibility of the recommendations which flow from issues she addresses.

Given Dr. Nummendal's concerns, it would appear that we should abandon our current CT assessment research and start over. First, we should search for a concept of CT that connects with real life, not just the college classroom. Second, we should scrap all existing CT assessment devices (the commonly used and the not so commonly used) and, once we finally learned what CT really is, we would start to develop "authentic" assessment instruments. Third, we should reconsider our assumptions about the proper curricular strategies and pedagogies to use in teaching CT so that the fruits of our assessments could inform and enrich college level CT instruction.[1]

---

[1] Dr. Nummendal makes other helpful points as well. Noting the research on cognitive and epistemic development, she advises that we be careful about what we expect by way of CT of 22 year olds. She calls for establishing baseline data to ground longitudinal research. And, she advises that we test for changes in both the level of performance and the variability of performance as students move through college.

Speaking to the issue of diversity as it applies to assessment, Dr. Nummendal recommends, "We need to consider multiple models for outcomes... we need to create measures that capture this diversity and to consider diversity as an asset, not a problem to fix." Alexander Astin's research on college outcomes indicates that diversity and a campus culture that openly addresses questions of diversity are both assets to academic growth and development. Students learn more in a context of social diversity.

But Dr. Nummendal's call for "multiple models for outcomes" and "measures to capture diversity" is a different thing. To say the least, it is extremely unclear. She might mean that out of fairness to the differently-abled we must devise equivalent, but alternative adaptations of the basic assessment program. If we do not, these persons will not have an equal chance to show their CT skills and dispositions. And, if this is what she means, then her recommendation makes sense.

But how does this call for multiple models apply to ethnic, religious, socioeconomic, cultural or gender diversity? Is the claim that we must alter how we assess persons who come from diverse groups? Or should we alter the questions, not the mode of assessment?

On the intuitive level, Dr. Nummendal is correct about the third recommendation; we do need to reconsider the typical pattern of instruction in college level CT. But not because we don't know what CT really is, nor how to assess it. Two different factors suggest this pedagogical review. First the evidence about the superior effectiveness of collaborative learning throughout the curriculum requires our attention.[2] Given the highly social contexts of problem solving, policy formation, and decision making (in and out of college life) a CT pedagogy grounded in cooperative, interactive learning promises to be far superior and more realistic than an individualistic pedagogy based on divisive competition.

Apart from the above, we also have empirical data about CT instruction. Dr. Nummendal's claim that "we still know relatively little about the effectiveness of various college [CT] courses" notwithstanding, the objective evidence suggests support for her call for pedagogical reform. Gender differences which have emerged on CT posttests, but which did not exist to a statistically significant degree on CT pretests or among control groups, suggest that the typical college level CT course differentially advantages men over women. Given the dozens of other variables controlled for in that research, and the current tendency to use competitive (argument pro and con, I win you lose) vs. collaborative (how can we solve this problem) pedagogy is clearly the variable which next needs to be investigated.[3]

These empirical findings suggest a next step in a systematic program of scientific inquiry precisely because Dr. Nummendal's first two recommendations have already been addressed successfully by researchers in CT assessment. There is reasonable accord about an appropriate and rich conceptualization of CT. We know how to conduct valid and reliable CT assessment, and have developed at least a few instruments suitable for such an analysis.[4]

The image of confused academicians hopelessly locked in their discipline-specific

---

Or should we alter our norms or our expectations of achievement? If so, why – what is the empirical or theoretical basis for demands such as these? Is it because persons from diverse groups can't be expected to achieve comparable levels of success because they hold different values, know different facts, or make different assumptions about the everyday world? And, if that is the concern, are we saying that it is impossible to devise assessment instruments which account for this? How far does this all go? In its most pernicious form, and clearly not a form suggested by Dr. Nummendal, this kind of proposal can be very elitist and socially harmful. We need less politics and far more science here.

2 A comprehensive, readable and sensible case for this is presented in the FIPSE funded "First Report of the Harvard Assessment Seminar" 1990, by Richard J. Light, Graduate School of Education, Kennedy School of Government, Cambridge, MA, 02138.

3 The evidence for a differential impact of standard CT instruction by gender is presented in "Technical Report #3, Gender, Ethnicity, Major, CT Self-Esteem, and the California Critical Thinking Skills Test" (ERIC Doc. No.: ED 326 584). Additional information about the California Critical Thinking Skills Test is available in "Using the CCTST in Research, Evaluation and Assessment," (ERIC Doc. No.: TM 017 349). This publication is also available from the California Academic Press, 217 La Cruz Ave., Millbrae, CA., 94030.

4 The research leading to a consensus among CT experts about what CT means is described below. The objective CT assessment tool referred to is the newly published "California Critical Thinking Skills Test: College Level," 1990, The California Academic Press, Millbrae, CA. 94030.

BEST COPY AVAILABLE

world views and failing to comprehend the true life meaning of CT is untrue. From before John Dewey, down through Max Black, to the present, leading figures in educational philosophy and public policy have argued for essentially the same thing Dr. Nummendal seeks -- that one of the central outcome of a college education must be critical thinking. It is a virtual article of faith among those committed to liberal education that CT, along with a breadth of knowledge, is the liberating force which frees students from the slavery of ignorance and frees them for productive citizenship, a fuller appreciation of the human condition, and lifelong learning.

But is this merely academic rhetoric, the stuff of college bulletins and mission statements? Or is there evidence that any of this really might be happening? Before deciding higher education in general, and CT education in particular, is a CT wasteland, let's see what the research says. Alexander W. Astin, of the UCLA Higher Education Research Institute, now completing a five year study of college graduates from 159 institutions, is a source of empirical information about both the objectively tested and the self-reported outcomes of an undergraduate education. Another important resource is the Penn State Center for the Study of Higher Education, directed by James L. Ratcliff, which includes the National Assessment Center.

What Astin is saying is that, based on the self-reports of graduates, their GRE, LSAT, MCAT, and NTE scores, and professional accomplishments, college students improve in their critical thinking through their studies in the general education program and by student-student and student-faculty interaction.

Well even so, maybe Dr. Nummendal is correct when she suggest that we are confused about all this because we really don't know what CT is.

However, we do. In 1990 a national panel of 46 CT experts, drawn from several different disciplines and kinds of colleges, completed two years of careful collaborative work with the publication of a well-argued, detailed consensus regarding the core college level critical thinking skills and dispositions. Each skill and sub-skill they identified qualifies not as a discipline specific factor, but a genuinely transferable cognitive skill that can be used in either a social or individual real life problem solving context. The same is true of the experts' consensus list of dispositions characteristic of a good critical thinker. Some CT dispositions relate to life and living in general, some to how one approaches problems and questions. The citations at the end of this review reveal a bit of the richness, discipline-neutrality, and practical utility behind this consensus conceptualization of CT. [5]

A word is in order about the discipline-specific vs. discipline-neutral dispute, which Dr. Nummendal says is "far from over." It is the general consensus among CT theorists and

---

[5] The consensus findings are published in Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (ERIC Doc. No.: ED 315 423), [80 pages including appendices]. Known as "The Delphi Report," this research was conducted under the auspices of the American Philosophical Association during 1988 and 1989 and published in early 1990. A 22 page "Executive Summary," which includes all tables, findings and recommendations of the Delphi Panel, is published by the California Academic Press, 217 La Cruz, Millbrae, CA. 94030.

assessment experts that this is a dead issue. Robert H. Ennis, perhaps the leading figure in CT and CT assessment in the nation, handily laid that matter to rest some years back. There may be some nicely refined twists put on some CT applications in advanced doctoral research done within given disciplines, but for the college level the list of core CT cognitive skills (analysis, interpretation, evaluation, inference, explanation, and self-regulation) are both practical and generic. The CT dispositions are as well. Professor Ennis directs the Illinois Critical Thinking Project out of the Department of Educational Policy at the University of Illinois at Urbana-Champaign. That unit has been conducting research on both everyday content and discipline-specific contexts for CT assessment for a number of years.

To become better connected with advances in theory development and empirical research in the area of CT, one might also contact The Center for Critical Thinking at Montclair State College in New Jersey, directed by Mark Weinstein; The Center for Critical Thinking and Moral Critique at Sonoma State University, directed by Richard Paul; and the Institute for Educational Research and Development at the University of Newfoundland, particularly Dr. Stephen Norris. Other persons with practical experience and technical, scholarly expertise in CT assessment who should be consulted include Joanne Carter-Wells, Dept. of Reading, CSU, Fullerton, and Barbara M. Lawrence, Coordinator of Student Outcomes Assessment at Idaho State University, and Marcia Mentkowski, of Assessment Office of Alverno College.

Dr. Nummendal raises a helpful, new issue for CT assessment when she calls for strategies suitable for assessing individuals as they work in social or collaborative contexts. A possible response to this takes us into the assessment of a different, but related skill/disposition, namely leadership. One model applicable to individual assessment in the context of group activity is that developed and used by the National Association of Secondary School Principles (NASSP) in their leadership assessment program. This technically sophisticated mode of assessment uses trained expert observers to note and record a wide variety of verbal and non-verbal cues which indicate leadership skills and dispositions in a variety of individual and group contexts. Although labor intensive and time consuming for both the assessors and the candidates, this process has withstood legal and theoretical challenges and become a standard mode for objectively evaluating adult leadership potential in this professional field. It might be adaptable for CT assessment. And with this possibility recognized, the CT assessment research agenda just expanded.

In summary, as tempting as it might first seem, we do not have to begin the science of CT assessment all over again. There have been false starts, and not everyone agrees on everything. But some positions are objectively more solid than others. Some research is superior. We have a rich, practical, and widely acceptable concept of CT which grew out of the application of qualitative social science methodology. We have devised and validated workable objective assessment devices. Using these, we have begun to learn important things about CT instruction. Achievements like these are leading us to new projects and new research questions.

Yes, the research agenda is far from exhausted. But the historically interesting issues

raised in Dr. Nummendal's paper, when compared to the extant research and successful projects, are insufficient to persuade us to abandon decades of work and start over. For scientific progress in CT assessment to continue we must pursue the research agendas now in place. We must build on our successes and respond to the challenges suggested by the objective data on college student CT abilities now starting to become available. We know what CT is and we have begun to find successful ways to assess it. Let's move ahead.

Consensus findings cited from Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (c) 1990, The California Academic Press, Millbrae, CA.

TABLE 1

CRITICAL THINKING AND THE IDEAL CRITICAL THINKER

"We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based. CT is essential as a tool of inquiry. As such, CT is a liberating force in education and a powerful resource in one's personal and civic life. While not synonymous with good thinking, CT is a pervasive and self-rectifying human phenomenon. The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. Thus, educating good critical thinkers means working toward this ideal. It combines developing CT skills with nurturing those dispositions which consistently yield useful insights and which are the basis of a rational and democratic society."

TABLE 3

CRITICAL THINKING COGNITIVE SKILLS AND SUB-SKILLS

1. Interpretation	• Categorization • Decoding Significance • Clarifying Meaning
2. Analysis	• Examining Ideas

	• Identifying Arguments
	• Analyzing Arguments
3. Evaluation	• Assessing Claims
	• Assessing Arguments
4. Inference	• Querying Evidence
	• Conjecturing Alternatives
	• Drawing Conclusions
5. Explanation	• Stating Results
	• Justifying Procedures
	• Presenting Arguments
6. Self-Regulation	• Self-Examination
	• Self-Correction

TABLE 5

AFFECTIVE DISPOSITIONS OF CRITICAL THINKING

APPROACHES TO LIFE AND LIVING IN GENERAL

- Inquisitiveness with regard to a wide range of issues
- Concern to become and remain generally well-informed
- Alertness to opportunities to use CT
- Trust in the processes of reasoned inquiry
- Self-confidence in one's own ability to reason
- Open-mindedness regarding divergent world views
- Flexibility in considering alternatives and opinions
- Understanding of the opinions of other people
- Fair-mindedness in appraising reasoning
- Honesty in facing one's own biases, prejudices, stereotypes, egocentric or sociocentric tendencies
- Prudence in suspending, making or altering judgments
- Willingness to reconsider and revise views where honest reflection suggests that change is warranted

APPROACHES TO SPECIFIC ISSUES, QUESTIONS, OR PROBLEMS

- Clarity in stating the question or concern
- Orderliness in working with complexity
- Diligence in seeking relevant information
- Reasonableness in selecting and applying criteria
- Care in focusing attention on the concern at hand
- Persistence through difficulties are encountered
- Precision to the degree permitted by the subject and the circumstances

Review of Nummedal's Paper "Designing a Process to Assess Higher-Order Thinking Skills in College Graduates: Issues of Concern"

by

Ronald K. Hambleton
University of Massachusetts at Amherst

Professor Nummedal addresses three important issues about assessing higher-order thinking skills in her paper: (1) domain of interest for the assessment, (2) selection of skills, and (3) assessment to inform practice. Her paper is thoughtful and insightful, and if her ideas are adopted, national assessments will be extremely difficult to carry out. I'm not suggesting that she is wrong in any way. I'm simply noting the implications of her interesting thoughts for the national assessment of Objective 5 of Goal 5.

Professor Nummedal argues that we should focus our attention in addressing Objective 5 of Goal 5 on "those skills that are acquired through learning experiences in college that are relevant to successful functioning in real-life situations" (p. 4). My problem with her focus is that it appears to de-emphasize skills that are important and oriented to academic outcomes. I believe at least some of the focus of any national assessment should be on the preparedness of graduates to use higher-order thinking skills in their chosen disciplines. Generic skills in critical thinking, reasoning, problem-solving, etc., in the context of academic pursuits, seem important to monitor and assess as well as the manifestation of these same skills in day-to-day activities.

Professor Nummedal seems to have a plan in mind where students develop their higher-order thinking skills in a college or university course. She writes about the skills that could be taught in such courses and how feedback might be given. I'm not opposed to her plan (how could anyone be opposed to a

course to develop critical thinking skills and provide students with feedback?). I can see the substantial merit of such a plan, but I was assuming an instructional and assessment plan in which students would (hopefully) develop higher-order thinking skills through their regular college or university coursework and then be assessed (or a sample of students would be assessed) in a "NAEP-like" way. Perhaps there is room in the master plan of national assessment to represent both paradigms.

Professor Nummedal focused in her paper on "every-day thinking," critical thinking courses, individual assessments, the use of "ill-structured problems" in assessment, team efforts to solve problems, etc. I agree very much with her ideas for improving aspects of instruction and assessment at the college level. I doubt, however, that a national assessment system could evolve from her work. Of course, now the dilemma is clear: Why teach critical thinking one way in the classroom and then have a national assessment system which is only a crude proxy for what is really important?

Professor Nummedal has presented measurement specialists with challenges that may not be solved for a long time in the context of a national assessment of critical thinking: The use of ill-structured problems, providing feedback during the assessment process (as might happen in practice), group problem-solving, scoring problems, questions about generalizability of performance scores to other contexts, problems of data aggregation across components of a task, across tasks, and even across students, etc. Professor Nummedal also challenges the system designers to provide results that will be useful at the <u>classroom level</u> as well as the national level.